# Databricks Lakehouse Platform : Administration, Développement et Optimisation

*Lien : https://innov-systems.com/formation/databricks-lakehouse-platform-administration-developpement-et-optimisation*

DURÉE
**5 jours (35h)**

RÉFÉRENCE
**BDI151**

CATÉGORIE
**Data Science et Data Engineering**

Innov Systems

## OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre l’architecture de la Databricks Lakehouse Platform
- Maîtriser l’administration des environnements Databricks
- Développer des pipelines de données performants
- Utiliser Delta Lake pour la gestion des données
- Optimiser les performances des workloads Spark
- Gérer la sécurité et la gouvernance des données
- Déployer des solutions data en production
- Intégrer Databricks avec les principaux clouds
- Concevoir des architectures data modernes
- Automatiser les traitements de données
- Exploiter les outils de data engineering et data science
- Étudier des cas concrets industriels

## POUR QUI ?

- Data engineers
- Data scientists
- Data analysts
- Cloud engineers
- Architectes data
- Développeurs BI
- DevOps engineers
- Consultants data
- Administrateurs systèmes cloud
- Chefs de projet data
- Toute personne impliquée dans les plateformes data modernes

Innov Systems

## Programme détaillé

### 1 / Introduction à la Databricks Lakehouse Platform

- Concept de Lakehouse (Data Lake + Data Warehouse)
- Architecture globale de Databricks
- Cas d’usage en data engineering et data science

### 2 / Architecture et composants principaux

- Workspaces, clusters et notebooks
- Delta Lake et stockage unifié
- Pipelines de données et orchestration

### 3 / Administration de la plateforme Databricks

- Gestion des utilisateurs et des droits (IAM)
- Configuration des clusters et ressources
- Sécurité et gouvernance des données

### 4 / Développement sur Databricks

- Utilisation des notebooks (Python, SQL, Scala)
- Développement de jobs et workflows
- Intégration avec les outils BI et data

### 5 / Delta Lake et gestion des données

- Transactions ACID sur les data lakes
- Versioning et time travel
- Optimisation des tables Delta

Innov Systems

### 6 / Data engineering avec Databricks

- Construction de pipelines ETL/ELT
- Traitement batch et streaming
- Orchestration des flux de données

### 7 / Performance et optimisation des workloads

- Optimisation des clusters Spark
- Gestion de la mémoire et du calcul distribué
- Réduction des coûts d’exécution

### 8 / Sécurité et gouvernance des données

- Unity Catalog et gestion des métadonnées
- Contrôle d’accès granulaire
- Conformité et audit des données

### 9 / Intégration avec l’écosystème cloud et BI

- Connexion à Azure, AWS et GCP
- Intégration avec Power BI, Tableau, etc.
- API et automatisation

### 10 / Études de cas et projets pratiques

- Construction d’un pipeline complet sur Databricks
- Analyse de données à grande échelle
- Bonnes pratiques de production

Inov Systems

## Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

## Prochaines dates programmées

| | |
|---|---|
| 12 au 18 Jan. 2027 | Présentiel - Casablanca |
| 08 au 12 Mars 2027 | Présentiel - Casablanca |
| 03 au 07 Mai 2027 | Présentiel - Casablanca |
| 28 Juin au 02 Juil. 2027 | Présentiel - Casablanca |
| 23 au 27 Août 2027 | Présentiel - Casablanca |

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

## Réservation & Renseignements

**Téléphone :** +212 522 247 210
**Email :** contact@innov-systems.com
**Web :** https//www.innov-systems.com


Innov Systems — Tous droits réservés